Published online February 6, 2007

Molecular Analysis of Soybean Lines with Low Palmitic Acid Content in the Seed Oil

Andrea J. Cardinal,* Joseph W. Burton, Ana Maria Camacho-Roger, Ji H. Yang, Richard F. Wilson, and Ralph E. Dewey

ABSTRACT

Palmitic acid is the major saturated fatty acid found in soybean [*Glycine max* (L.) Merr.] oil, accounting for approximately 11% of the seed oil content. Reducing the palmitic acid levels of the oil is desirable because of the negative health effects specifically associated with this fatty acid. One of the genetic loci known to reduce the seed palmitate content in soybean is *fap$_{nc}$*. Previous studies indicated that *fap$_{nc}$* is associated with a deletion in a gene (designated *FATB*) encoding 16:0-ACP thioesterase activity. In this report, we isolated full length cDNAs of three of the four unique *FATB* genes expressed in soybean and show that the isoform designated *GmFATB1a* represents the specific gene deleted in lines possessing the *fap$_{nc}$* locus. Allele specific primers corresponding to *GmFATB1a* were used to genotype plants from two F4-derived populations that were segregating for *fap$_{nc}$*. The *GmFATB1a*-specific markers were effective in accounting for 62 to 70% of the genotypic variation in palmitate content in the two populations studied. Because the markers developed in this study are 100% linked to the locus of interest, they should be particularly useful in marker-assisted selection programs designed to lower the palmitic acid levels of soybean oil.

SOYBEAN is an important crop that provides 57.7% of the world's oilseed production (USDA statistics, 2003). Soybean oil is mostly composed of triacylglycerols (TAGs). The relative composition of saturated and unsaturated fatty acids in the seed TAGs determines the quality of an oil. Palmitic acid (16:0) is the predominant saturated fatty acid in soybean oil. Oil from typical soybean cultivars contains 110 to 120 g kg^{-1} palmitic acid (Burton et al., 1994; USDA-ARS, National Genetics Resources Program). Consumption of oils with reduced palmitic acid content is desirable to reduce the health risks of coronary diseases and breast, colon, and prostate cancer properties associated with this fatty acid (Hu et al., 1997; Henderson, 1991).

Soybean oil quality can be improved through genetic alteration of the plant. Several lines with reduced palmitic acid concentration have been developed by mutagenesis. C1726 (85.4 g kg^{-1} 16:0) and ELLP2 (71.4 g kg^{-1} 16:0) were developed by ethyl methanesulfonate (EMS) mutagenesis of the cultivars Century and Elgin87, respectively. A22 (78.4 g kg^{-1} 16:0) was developed by *N*-nitroso-*N*-methyl urea mutagenesis of A1937, and J3 (57.4 g kg^{-1} 16:0) was developed from X-radiation of cultivar Bay (Fehr et al., 1991a; Stojšin et al., 1998; Wilcox and Cavins, 1990; Rahman et al., 1996). In addition, one natural mutation was discovered in line N79-2077 (64 g kg^{-1} 16:0) and its derived line N79-2077-12 from the fifth cycle of recurrent selection for high oleic concentration (Burton et al., 1983, 1994). The alleles conferring low palmitic acid in these lines were given different names: *fap*1 (in C1726), *fap*3 (in A22), *fap$_{nc}$* (in N79-2077), *fap** or *fap*x (in ELLP2), and *sop*1 (in J3) (Schnebly et al., 1994; Stojšin et al., 1998; Wilcox and Cavins, 1990; Kinoshita et al., 1998; Primomo et al., 2002). Of those five alleles, three are at independent loci from each other (*fap1*, *fap3*, and *fap**) (Schnebly et al., 1994; Stojšin et al., 1998; Wilcox et al., 1994; Kinoshita et al., 1998; Primomo, 2000; Primomo et al., 2002). *fap$_{nc}$* is allelic to *fap3* (Primomo, 2000; Primomo et al., 2002). *sop1* is independent of the *fap1* locus, but its allelic relationship to the other loci is unknown (Kinoshita et al., 1998).

In addition to the major genes mentioned above, minor-modifier genes influence the low palmitic acid content in soybean oil (Stojšin et al., 1998; Rebetzke et al., 1998b, 2001; Li et al., 2002; Horejsi et al., 1994). Rebetzke et al. (1998b) observed heritability estimates for 16:0 content on a plot basis between 0.37 and 0.73 and on an entry-mean basis between 0.83 and 0.96. They concluded that selection based on evaluations in a few environments for both major and minor genes would be successful.

Correlated changes in the other saturated and unsaturated fatty acids in the soybean oil have been observed in low palmitic acid lines that carry the major and minor genes described above. For example, when lines homozygous for the *fap$_1$* and *fap$_3$* alleles (<41.5 g kg^{-1} 16:0) were compared with normal lines (>100 g kg^{-1} 16:0) derived from the same population, the low palmitic acid lines tended to have reduced stearic acid (18:0) and increased oleic acid (18:1), linoleic acid (18:2), and linolenic acid (18:3) contents (Ndzana et al., 1994). Similarly, progenies that inherited the major *fap$_{nc}$* allele from the cross of N87-2122-4 to two high yielding normal palmitate lines had a 6 to 15% reduction in stearic acid, (Rebetzke et al., 1998b, 2001) a 4 to 10% increase in oleic acid content in both crosses (Rebetzke et al., 1998a, 2001) and an increase in linolenic acid content in one cross (Rebetzke et al., 1998a). No differences in linoleic acid were observed between progenies that inherited the major *fap$_{nc}$* allele and those that inherited the normal allele (Rebetzke et al., 1998a). However,

Andrea J. Cardinal, Ana Maria Camacho-Roger, Ji H. Yang, and Ralph E. Dewey, Dep. of Crop Science, North Carolina State Univ., Raleigh, NC 27695; Joseph W. Burton and Richard F. Wilson, USDA-ARS, Plant Sci. Res., 3127 Ligon St., Raleigh, NC 27607. This project was funded by the United Soybean Board. Received 2 May 2006. *Corresponding author (andrea_cardinal@ncsu.edu).

Published in Crop Sci. 47:304–310 (2007).
Genomics, Molecular Genetics & Biotechnology
doi:10.2135/cropsci2006.04.0272

677 S. Segoe Rd., Madison, WI 53711 USA

Abbreviations: BLUP, best linear unbiased predictor; EST, expressed sequence tags; FATA, 18:1-ACP thioesterase gene; FATB, 16:0-ACP thioesterase gene; KAS II, 3-ketoacyl-ACP synthetase II; PCR, polymerase chain reaction; TAGs, triacylglycerols; 16:0, palmitic acid; 18:0, stearic acid; 18:1, oleic acid; 18:2, linoleic acid; 18:3, linolenic acid.

there was significant genetic variation within each major fap_{nc} group because of the segregation of minor genes (Rebetzke et al., 1998a). Within these groups, there was a significant negative genetic correlation between palmitate and oleate and a significant positive correlation between palmitate and linolenate (Rebetzke et al., 1998a). Palmitate content in soybean oil was negatively correlated with linoleate content and positively correlated with oleate content in a population that was segregating for the fap_1 and *fan* loci (Nickell et al., 1991).

Several steps in the TAG biosynthetic pathway could affect the palmitate concentration. In the case of low palmitic acid mutants, changes in activity or substrate specificity in 3-ketoacyl-ACP synthetase II (KAS II), 16:0-ACP thioesterase, lysophosphatidic acid acyltransferase, glycerol-3-phosphate acyltransferase, or diacylglycerol acyl transferase enzymes could be responsible for their phenotype. A biochemical study comparing C1726 (*fap*1*fap*1), N79-2077-12 ($fap_{nc}fap_{nc}$) and 'Dare' found no evidence that glycerolipid acyltransferase activity or the biosynthetic steps after the formation of acyl-CoA were altered in these lines. The TAG composition of these lines was related to the amount of 16:0 produced in the plastids (Wilson et al., 2001b, 2001c). Northern analysis revealed that $fap_{nc}fap_{nc}$ genotypes have reduced accumulation of transcripts corresponding to a 16:0-ACP thioesterase (FATB) gene and no difference in transcript accumulation for genes encoding 18:1-ACP thioesterase (FATA) or 18:0-ACP desaturase activities (Wilson et al., 2001b). No differences in transcript accumulation for *fap*1*fap*1 genotypes were observed for any of the genes assayed (Wilson et al., 2001b). Southern analysis using the FATB cDNA as a probe revealed that multiple copies of the 16:0-ACP thioesterase gene are present in the soybean genome and that the $fap_{nc}fap_{nc}$ genotype has a deletion in one of those copies (Wilson et al., 2001a). It is not known which of the FATB gene copies is deleted in the $fap_{nc}fap_{nc}$ genotype, where in the soybean genome these FATB genes are located, and how much of the palmitate phenotypic variation each isoform accounts for.

The objectives of this study are to (i) identify the specific *FATB* gene that is deleted in low palmitic acid soybean lines homozygous for the fap_{nc} mutation, (ii) develop allele specific PCR markers corresponding to this gene, and (iii) estimate the amount of phenotypic variation in the 16:0, 18:0, 18:1, 18:2, and 18:3 composition that is explained by this locus in two adapted populations grown in replicated trials.

MATERIALS AND METHODS

Amplification and Sequence Analysis of Full Length Soybean FATB cDNAs

Total RNA was isolated from developing seeds of soybean cultivar Century using the TRIzol reagent according the manufacturer's protocol (Invitrogen, Carlsbad, CA). Poly (A)+ RNA was subsequently recovered from total RNA using the Messagemaker system (Invitrogen). First-strand cDNA was synthesized from poly (A)+ RNA using an oligo-dT primer and SuperScript II reverse transcriptase (Invitrogen).

Sequence information from soybean ESTs available in GenBank was used to synthesize PCR primers corresponding to the 3′ flanking regions of the *GmFATB1a*, *GmFATB1b*, and *GmFATB2a* genes. To obtain sequence information from the respective 5′ flanking regions, reverse primers directed against individual *GmFATB* genes were used together with a vector-specific primer and a soybean cDNA library that we had previously cloned into a yeast expression vector as a template (Dewey et al., 1994). The following primers were designed to amplify full length cDNAs using the above described first-strand cDNA from Century as the template: forward primer 5′-TCTACCGGAGAAGCGACCT-3′ and reverse primer 5′-CCCCCTTACCCACCAAAGTA-3′ (*GmFATB1a*); forward primer 5′-TACCCTACAAGCTGCCGTAG-3′ and reverse primer 5′-CCCCCTTACCCACCAAAGTA-3′ (*GmFATB1b*); and forward primer 5′-GTAAGGCGTAGGGGGTTGTT-3′ and reverse primer 5′-ACACCAAGCAGCTGCATAAC-3′ (*GmFATB2a*). Nucleic acid and predicted protein sequences were aligned and compared using the CLUSTALW (Thompson et al., 1994) and GAP (University of Wisconsin Genetics Computing Group) algorithms.

Amplification of Genomic DNAs

Genomic DNA was prepared from young soybean leaf tissue as previously described (Murray and Thompson, 1980). Primers 5′-GACATAGTTCAAGTGGACACT-3′ (forward) and 5′-TTCACAACACCAAAGTTGTTCAC-3′ (reverse) were used to amplify a 1178-bp fragment of *GmFATB1a* genomic DNA. The corresponding region of *GmFATB1b* genomic DNA was amplified using primers 5′-GACGTAGTTCAAGTAGACACC-3′ (forward) and 5′-TTCACAACACCAATGTTGTTCAT-3′ (reverse). Three introns were observed within this specific region of the *GmFATB1* genes. The partial genomic sequence information generated by these primers is shown in supplemental Fig. S1 and have been deposited in GenBank as Accession no. DQ861997 (*GmFATB1a*) and no. DQ861998 (*GmFATB1b*).

GmFATB1a- and *GmFATB1b*-Specific Molecular Markers

On the basis of the results of a *GmFATB1a* versus *GmFATB1b* genomic sequence analysis, PCR primers were designed against regions predicted to have sufficient polymorphism to yield gene-specific amplification products. A 312-bp fragment of *GmFATB1a* (marker FATB1a-312) was amplified by primers 5′-TGACATAGTTCAAGTGGACACT-3′ (forward) and 5′-GCAAAATGCAAATGATTACCTG-3′ (reverse). The corresponding region of *GmFATB1b* (marker FATB1b-314) was amplified by primers 5′-GACGTAGTTCAAGTAGACACC-3′ (forward) and 5′-GCAAAATTCAAATGATTTCCTC-3′ (reverse). Specificity of the primers was verified by sequence analysis of the fragments amplified by each primer pair. The sequences of the amplified products corresponded to the expected sequences of the *GmFATB1a* and *GmFATB1b* fragments (data not shown). Similarly, an additional marker pair was generated to a separate region of the *GmFATB1* genes using the following primer pairs: 5′-GCATCTTCATTTTGCAGC-3′ (forward) and 5′-TTCACAACACCAAAGTTGTTCAC-3′ (reverse) (*GmFATB1a*-specific marker FATB1a-411); and 5′-TTATGTATCTTCATTCTTCCAGT-3′ (forward) and 5′-TTCACAACACCAATGTTGTTCAT-3′ (reverse) (*GmFATB1b*-specific marker FATB1b-417). PCR reactions were conducted in a standard buffer (10 m*M* Tris-HCl, pH 8.3, 50 m*M* KCl, 1.5 m*M* $MgCl_2$) with 200 μM dNTPs, 20 pmol each primer, 50 ng of genomic DNA template and

2.5 U Taq polymerase in a final reaction volume of 50 μL. Thermocycling was conducted with an initial 5 min incubation at 94°C, followed by 30 cycles of 94°C for 30 s, 52°C for 30 s, 72°C for 1 min, and a final 4 min extension at 72°C. Amplification products were visualized following separation on 1.2% (w/v) agarose gels containing 0.3 μg mL^{-1} ethidium bromide.

Development of Populations

Two F4-derived populations were developed by single seed descent (Brim, 1966). One population (FAHH00) was derived from the cross of a high yielding line, 'Corsica', and a low linolenic, low palmitic acid line, N97-3681-11. A second population (FADD00) was derived from the cross of high yielding line, 'Brim', and a low linolenic, low palmitic line, N97-3708-13. N97-3681-11, N97-3708-13 and the cultivar Satelite are sister lines that were derived from different F3 plant selections from the cross 'Soyola' × {'Brim'(2) × [N88-431(2) × (N90-2013 × C1726)]}. Line N88-431 was selected from the cross N88-1299 × N82-2037. N84-1299 was a selection from the first cycle of a high protein recurrent selection population (Holbrook et al., 1989). N82-2037 had Gasoy17 as its maternal parent and N77-940 as the paternal parent. C1726 is a reduced palmitic acid selection from EMS treated Century and is homozygous for the fap_1 mutation (Wilcox and Cavins, 1990). N90-2013 is a reduced palmitic acid selection from the cross PI123,440 × N79-2077-12 (Burton et al., 1994). Soyola is a low linolenic acid line (Burton et al., 2004). N97-3681-11 and N97-3708-13 are F4-derived lines and are assumed homozygous for the low palmitic fap_{nc} mutant allele from N79-2077-12, the low palmitic fap_1 mutant allele from C1726, and the low linolenic *fan*(PI123440) allele. In addition, N97-3681-11 and N97-3708-13 are highly related to Brim, and the coefficient of parentage between these lines and Brim is 0.777 (data not shown). Therefore, in the Brim × N97-3706-13 population only 22.3% of the soybean genome is expected to be segregating.

Phenotypic Evaluation of Populations

Each population was evaluated in separate field experiments. Ninety-eight F4-derived lines and the parents of the FADD00 population and 99 F4-derived lines and one parent (N97-3681-11) of the FAHH00 population were grown at Plymouth, NC, in 2002 and Caswell and Plymouth, NC, in 2003. Within each environment, the experimental design for each population was a 10 × 10 lattice with two replications. At harvest time, a 30 g subsample of seed from each plot was analyzed for fatty acid content, by gas-liquid chromatography of the methyl esters, at the USDA-ARS laboratory, Raleigh, NC. Fatty acid content is reported in grams per kilogram of total lipids.

Genotypic Evaluation of Populations

DNA was extracted by a CTAB (hexadecyltrimethylammonium bromide) protocol from a bulk of leaf tissue from several plants of each line (Li et al., 2002). FATB1a-312 and FATB1a-411 allele specific primer markers were used to perform PCR analysis with genomic DNA from each line in the two populations. The PCR reactions consisted of: 0.8 μL of 5 pmol/μL forward and reverse primers, 0.5 μL of 3.125 m*M* dNTPs, 3 μL of 15 m*M* MgCl, 0.05 μL of Taq polymerase, 3.35 μL of ddH20, and 5 μL of 10ng/μL DNA. The PCR protocol included a first step of denaturation at 95°C for 5 min, then a second step of denaturation at 92°C for 1 min, a third step of annealing at 59°C for 1 min, a fourth step of extension at 69°C for 2 min, then steps two to four were subsequently repeated for 36 cycles. The reactions were performed in a 96-well PTC-100 thermal cycler (MJ Research, Inc., Waltham, MA) or a 384-well PTC-200 thermal cycler (MJ Research, Inc.). The amplified fragments were separated in 4% (w/v) SFR agarose (AMRESCO, Inc., Solon, OH) gels stained with ethidium bromide and visualized under UV light. The genotype of the FATB1a-312 and FATB1a-411 markers were scored as a presence or absence of a 312- or 411-bp band.

Statistical Analysis of Populations

The fatty acid data from each population were analyzed as a lattice design across environments by PROC Mixed, SAS 8.2 (SAS Institute, Inc., 1999). Locations, replications, incomplete blocks, and lines were considered random effects. Since lines were considered a random effect and a mixed model approach was used, then a best linear unbiased predictor (BLUP) for each line was obtained by adding the overall mean effect to the random effect of each line (SAS Institute Inc., 1996). The BLUPs for each fatty acid for each line were then used to perform a *t* test to compare the effect of inheriting one or two copies of the FATB1a-312 or FATB1a-411 marker alleles (heterozygous or homozygous wild type) versus zero copies (homozygous mutant). The FATB1a-312 or FATB1a-411 marker class was the independent variable, and the BLUPs for each line for each fatty acid were the dependent variable in the analysis. The mean of lines homozygous or heterozygous for the FATB1a-312 or FATB1a-411 allele (presence of wild-type allele), the mean of lines homozygous for the absence of the FATB1a-312 or FATB1a-411 allele (mutant allele), the standard deviation of those means, the difference between means, and the *R*-square for the model were obtained using SAS Proc GLM, SAS 8.2 (SAS Institute, Inc., 1999).

RESULTS AND DISCUSSION

Isolation of Soybean FATB Genes

We have previously shown through RNA blotting assays that soybean lines homozygous for the fap_{nc} locus accumulate less FATB transcripts than normal soybeans (Wilson et al., 2001b). Subsequently, Southern blotting experiments showed that the fap_{nc} locus was associated with a FATB deletion event (Wilson et al., 2001a). Analysis of the numerous EST sequences with FATB homology that have been deposited in GenBank enabled us to define four highly similar yet distinct FATB sequences, designated *GmFATB1a*, *GmFATB1b*, *GmFATB2a*, and *GmFATB2b*. ESTs corresponding to *GmFATB1a* are the most prevalent, represented by 22 independent entries; ESTs corresponding to *GmFATB1b*, *GmFATB2a*, and *GmFATB2b* are found one, five, and seven times, respectively (data not shown).

Using a modified 5′-RACE strategy, full-length cDNAs of *GmFATB1a*, *GmFATB1b*, and *GmFATB2a* were amplified from the soybean cultivar Century and sequenced. The nucleotide and predicted protein products of *GmFATB1a* and *GmFATB1b* both share over 96% sequence identity. In contrast, approximately 73% nucleotide sequence identity (79% predicted protein identity) is shared between either *GmFATB1a* or *GmFATB1b* and the *GmFATB2a* sequence. Although full-length sequence information for *GmFATB2b* is lacking, alignments of the partial sequences derived from GenBank suggest that this sequence is approximately 95% identical to *GmFATB2a* (data not shown). These

results, the existence of four isoforms that can be assigned into two pairs by sequence similarity, are very similar to those reported for the omega-3 desaturase (*FAD3* gene) that converts linoleoyl-PC to produce linolenoyl-PC in soybean (Bilyeu et al., 2003; Anai et al., 2005). Four similar *FAD3* sequences were observed that were grouped as two closely related pairs on the basis of sequence similarity. Both the *FATB* and *FAD3* gene sequence observations are congruent with the hypothesis that the soybean genome had undergone both an ancient and a more recent duplication event (Shoemaker et al., 1996; Schlueter et al., 2004; Pfeil et al., 2005). According to this model, some soybean genes would be expected to be represented by four copies within the genome that could be separated in two pairs that would be more similar within each pair that originated from the more recent duplication event.

In silico analysis of the EST database indicated that the *GmFATB1a* gene isoform is the most prevalent. In addition, the fap_{nc} allele had the strongest low palmitic acid phenotype of all mutations characterized to date. From these observations, we speculated that *GmFATB1a* would be the best candidate for being the specific *FATB* isoform deleted in fap_{nc} lines. To test this possibility, we attempted to generate gene-specific PCR primers that would uniquely amplify the *GmFATB1a* isoform. However, because of the high sequence identity shared between *GmFATB1a* and *GmFATB1b*, we had difficulty finding sufficient polymorphisms for producing primers that would enable reliable, unambiguous amplification of these highly homologous genes. To overcome this problem we amplified a region of *GmFATB1a* and *GmFATB1b* genomic DNA (Fig. S1). By designing PCR primers on the basis of polymorphic regions found within both exon and intron regions, gene-specific markers capable of distinguishing *GmFATB1a* and *GmFATB1b* were created. As shown in supplemental Fig. S1, marker FATB1a-312 is defined by primers that specifically amplify a 312-bp fragment of the *GmFATB1a* gene when genomic DNA is used as a template; marker FATB1a-411 is generated using a primer pair that amplifies a 411-bp genomic fragment of *GmFATB1a*. Similarly, markers FATB1b-314 and FATB1b-417 are the products of *GmFATB1b*-specific primers that amplify 314- and 417-bp products, respectively.

Expected amplification products were observed for all four *FATB1* markers when genomic DNA from normal soybean cultivar Dare was used as a template (Fig. 1A). Soybean lines C1726 and C1727, which possess the *fap1* and *fap2* mutations, respectively, that alter the palmitic acid content of the soybean seed, yet are nonallelic to fap_{nc}, also yielded the expected PCR products for all four markers. In contrast, when genomic DNAs from two lines that are homozygous for the fap_{nc} locus (N79-2077 and N93-2008) were used as template, only markers corresponding to *GmFATB1b* amplified a DNA segment (Fig. 1A). These results indicate that *GmFATB1a* is the specific *FATB* isoform that is deleted in soybean lines possessing the fap_{nc} locus. Furthermore, PCR reactions using genomic DNA from the fap_{nc} containing low palmitate lines N97-3681-11 and N97-3708-13 and the *GmFATB1a* specific markers also failed to yield amplification products (data not shown). Cumulatively, our results strongly suggest that the molecular basis for the low palmitic acid fap_{nc} mutation is a deletion of all or a portion of the *GmFATB1a* gene.

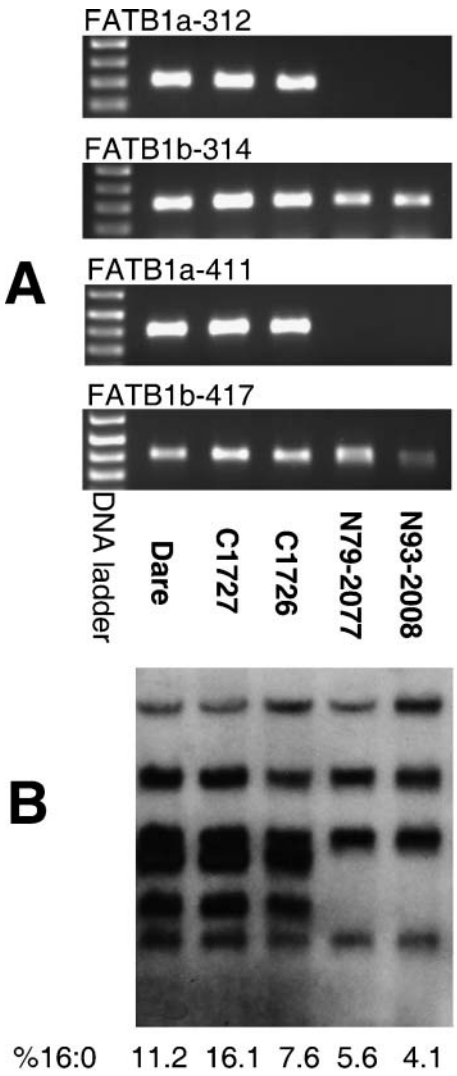

Fig. 1. Marker analysis of soybean lines possessing novel palmitic acid phenotypes. (A) PCR amplification products using *FATB1a*- and *FATB1b*-specific primers. Lines N79-2077 and N93-2008 are homozygous for the fap_{nc} locus. (B) Southern blot analysis of *Xba*I-digested genomic DNA of the genotypes shown in (A) using *GmFATB1a* as the hybridization probe. Palmitic acid content for each line is indicated.

Genotypic Evaluation of Populations

If deletion of the *GmFATB1a* gene represents the underlying basis of the fap_{nc} mutant allele, a strong correlation would be expected between the low palmitate phenotype mediated by fap_{nc} and the deletion of *GmFATB1a*. To test this prediction, the FATB1a-312 and/or FATB1a-411 specific markers were used to genotype two F4-derived populations segregating for fap_{nc} (see Materials and Methods). Individuals from segregating populations were initially screened using the FATB1a-312 marker. Samples that failed to yield com-

pletely definitive results were subsequently assayed with the FATB1a-411 marker to ensure accurate genotyping. DNA bulks of several individuals from each of 98 F4-derived lines of population FAD00 were assayed using the *GmFATB1a* specific markers. Of the 98 lines tested, seven cases were observed in which the fatty acid composition phenotype of a line did not match the fap_{nc} genotype. In these cases, the DNA from several individual plants of each line was newly extracted and the fap_{nc} genotype was obtained for each plant to test for line contamination or fap_{nc} segregation. A binomial test was used to test if the observed number of plants with Fap_{nc}- genotype in each F5:7 line was consistent with the hypothesis that the fap_{nc} was segregating in the line. If the hypothesis was rejected, then the line was declared to be contaminated. In six cases out of seven, the lines were shown to be contaminated. In the remaining one case, the line was clearly segregating for fap_{nc}. Of the 99 lines similarly assayed in population FAHH00, five cases were observed in which the phenotype of a line did not match the fap_{nc} genotype. In three cases, the lines were shown to be contaminated, and in the remaining two cases the lines proved to be segregating for the fap_{nc} allele. Therefore, in addition to following the low palmitic acid trait, use of the *GmFATB1a* specific markers also proved to be effective in detecting line contamination at a 5 to 10% level.

The number of observed homozygous lines for the fap_{nc} allele was significantly less than expected for F4-derived populations. A Chi-square test (not shown) demonstrated significant segregation distortion for this locus in both populations. No significant difference in maturity was observed between fap_{nc} fap_{nc} and Fap_{nc}– palmitate lines (data not shown); therefore, unintentional selection for later maturing lines during the inbreeding procedure cannot account for the segregation distortion. Only F4-derived lines that produced enough seed to be grown in replicated yield trials were used in this study. Therefore, if $fap_{nc}fap_{nc}$ lines tend to yield less than normal $Fap_{nc}Fap_{nc}$ palmitate lines, as was observed by Rebetzke et al. (1998b, 2001), selection against low yielding lines may have been responsible for recovery of fewer than expected $fap_{nc}fap_{nc}$ lines.

Phenotypic Evaluation of Populations

The mean palmitic acid content of seed oil from lines homozygous for the fap_{nc} allele is 48.2 g kg^{-1} in the Corsica × N97-3681-11 population and 54.3 g kg^{-1} in the Brim × N97-3708-13 population (Table 1). These means are lower than the expected from lines homozygous only for the $fap_{nc,}$ such as N79-2077 (64 g kg^{-1} 16:0) or its derived line N79-2077-12, because of the independent segregation of the additional low palmitate locus fap_1 in these populations (Burton et al., 1983, 1994). The mean palmitic acid content of lines heterozygous or homozygous for the presence of the FATB1a-312 or FATB1a-411 markers was 89.2 and 89.5 g kg^{-1}, in the two populations, respectively. These means are lower than the mean of Brim (110 g kg^{-1} of 16:0) because in addition to the segregating fap_1 locus, they also represent both the heterozygous and homozygous classes for the wild-type Fap_{nc} allele. The differences between the two groups were highly significant, and they explained 70.4 and 61.8% of the genetic variation in palmitic acid oil seed content in the two populations, respectively. The amount of genetic variation in palmitate content explained by this locus in these two populations is twice the amount explained by the major QTL detected in an F2 derived population segregating for the fap_{nc} allele only (Li et al., 2002).

Lines homozygous for the fap_{nc} allele also had significantly lower stearic acid content and significantly higher linoleic acid content in the seed oil in both populations (Table 1). Reduction in 18:0 suggests that the *GmFATB1a* encoded thioesterase also possesses some activity toward 18:0-ACP substrates, a result consistent with studies of 16:0-ACP thioesterase enzymes in other plant species (Jones et al., 1995; Voelker, 1996). The reduction in stearate content in lines homozygous for

Table 1. Palmitic acid, stearic acid, oleic acid, linoleic acid, and linolenic acid content means of seed oil of lines homozygous $fap_{nc}fap_{nc}$ and of lines heterozygous or homozygous (Fap_{nc}–) for the FATB1a allele, standard deviation of those means, difference between means and their significance, and the R^2 for the difference of those means for two F4-derived populations.

Corsica × N97-3681-11						
Fap_{nc} genotype	N†	Palmitic acid	Stearic acid	Oleic acid	Linoleic acid	Linolenic acid
		g kg^{-1}				
$fap_{nc}fap_{nc}$	22	48.2 (8.8)	28.0 (2.1)	291.2 (26.2)	572.8 (29.1)	54.8 (13.0)
Fap_{nc}	78	89.2 (11.7)	30.9 (2.3)	292.7 (29.3)	535.8 (26.5)	52.8 (11.7)
Difference		−41.0****	−2.9****	NS	37.0****	NS
R^2 (%)		70.4	22.7	0	24.6	0.5
Brim × N97-3708-13						
Fap_{nc} genotype	N	Palmitic acid	Stearic acid	Oleic acid	Linoleic acid	Linolenic acid
		g kg^{-1}				
$fap_{nc}fap_{nc}$	25	54.3 (8.8)	30.5 (1.4)	311.4 (18.1)	543.6 (27.1)	56.4 (11.3)
Fap_{nc}–	75	89.5 (13.0)	33.1 (1.0)	297.0 (21.6)	525.6 (27.6)	56.1 (11.5)
Difference		−35.2****	−2.5****	14.4**	18.0**	NS
R^2 (%)		61.8	50.4	8.5	7.6	0

** significant $p < 0.01$.
*** significant $p < 0.001$.
**** significant at $p < 0.0001$.
† N = number of lines with a specific Fap_{nc} genotype.

the *fap$_{nc}$* was also observed in other studies (Rebetzke et al., 1998b, 2001).

The influence of the *fap$_{nc}$* allele on oleic acid content was less clear. Lines homozygous for the *fap$_{nc}$* allele had significantly higher oleic acid content in the Brim × N97-3708-13 population, but not in the Corsica × N97-3681-11 population. Increased oleic acid content has previously been described in populations segregating for the *fap$_{nc}$* allele (Rebetzke et al., 1998a, 2001). The inheritance of the *fap$_{nc}$* allele had no effect on the linolenic acid content in either population (Table 1). This result is contrary to that observed by Ndzana et al. (1994) and Rebetzke et al. (1998a), but the populations studied in this project differ from those previous reports in that our populations are also segregating for the low linolenic acid *fan*(PI123440) allele. Previous studies have shown that the major low palmitic and low linolenic acid loci are independent from each other (Cherrak et al., 2003). It is possible that the populations used in previous studies by Ndzana et al. (1994) and Rebetzke et al. (1998a) were segregating for "normal" alleles that affect linolenate content but that are not allelic to *fan* so the correlation between palmitate and linolenate would differ in those populations.

Although the *fap$_{nc}$* allele affects the composition of several fatty acids in the seed oil, the amount of variation explained differed considerably between populations (Table 1). For example, the *fap$_{nc}$* allele accounts for 50.4% of the genetic variation for stearic acid content in the Brim × N97-3708-13 population but only 22.7% in the Corsica × N97-3681-11 population. Similarly, the *fap$_{nc}$* allele accounts for 7.6% of the linoleic acid genetic variation in the Brim × N97-3708-13 population and 24.6% in the Corsica × N97-3681-11 population.

In summary, we have determined that lines homozygous for the *fap$_{nc}$* mutation have a deletion in the *GmFATB1a* gene that encodes a 16:0-ACP thioesterase. The *GmFATB1a* gene product also influences 18:0 production because lines that have a deletion in this gene have reduced palmitate and stearate in their oil. The most likely explanation of this observation is that the *GmFATB1a* encoded enzyme also has some activity on 18:0-ACP substrates. Segregation at the *fap$_{nc}$* locus explained 62 to 70% of the genotypic variation in palmitate content in two F4-derived populations using specific markers we developed that uniquely amplify the *GmFATB1a* isoform. Future research goals will be to map the *GmFATB1a* gene and compare its genomic location with the major QTL for palmitate observed by Li et al. (2002).

REFERENCES

Anai, T., T. Yamada, T. Kinoshita, S.M. Rahman, and Y. Takagi. 2005. Identification of corresponding genes for three low-α-linolenic acid mutants and elucidation of their contribution to fatty acid biosynthesis in soybean seed. Plant Sci. 168:1615–1623.

Bilyeu, K.D., L. Palavalli, D.A. Sleper, and P.R. Beuselinck. 2003. Three microsomal omega-3 fatty-acid desaturase genes contribute to soybean linolenic acid levels. Crop Sci. 43:1833–1838.

Brim, C.A. 1966. A modified pedigree method of selection in soybeans. Crop Sci. 6:220.

Burton, J.W., R.F. Wilson, and C.A. Brim. 1983. Recurrent selection in soybean IV. Selection for increased oleic acid percentage in seed oil. Crop Sci. 23:744–747.

Burton, J.W., R.F. Wilson, and C.A. Brim. 1994. Registration of N79-2077-12 and N87-2122-4, two soybean germplasm lines with reduced palmitic acid in seed oil. Crop Sci. 34:313.

Burton, J.W., R.F. Wilson, W. Novitzky, and T.E. Carter, Jr. 2004. Registration of 'Soyola' soybean. Crop Sci. 44:687.

Cherrak, C.M., V.R. Pantalone, E.J. Meyer, D.L. Ellis, S.L. Melton, D.R. West, and J.R. Mount. 2003. Low-palmitic, low-linolenic soybean development. J. Am. Oil Chem. Soc. 80:539–543.

Dewey, R.E., R.F. Wilson, W.P. Novitzky, and J.H. Goode. 1994. The *AAPT1* gene of soybean complements a cholinephosphotransferase-deficient mutant of yeast. Plant Cell 6:1495–1507.

Fehr, W.R., G.A. Welke, E.G. Hammond, D.N. Duvick, and S.R. Cianzio. 1991. Inheritance of reduced palmitic acid content in seed oil of soybean. Crop Sci. 31:88–89.

Henderson, M.M. 1991. Correlations between fatty acid intake and cancer incidence. p. 136. *In* G. J. Nelson (ed.) Health effects of dietary fatty acids. AOCS, Champaign, IL.

Holbrook, C.C., J.W. Burton, and T.E. Carter. 1989. Evaluation of recurrent restricted index selection for increasing yield while holding seed protein constant in soybean. Crop Sci. 29:324–329.

Horejsi, T.F., W.R. Fehr, G.A. Welke, D.N. Duvick, E.G. Hammond, and S.R. Cianzio. 1994. Genetic control of reduced palmitate content in soybean. Crop Sci. 34:331–334.

Hu, F.B., Stampfer M. J., J.E. Manson, E. Rimm, G.A. Colditz, B.A. Rosner, C.H. Hennekens, and W.C. Willett. 1997. Dietary fat intake and the risk of coronary heart disease in women. N. Engl. J. Med. 337:1491–1499.

Jones, A., H.M. Davies, and T.A. Voelker. 1995. Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases. Plant Cell 7:359–371.

Kinoshita, T., S.M. Rahman, T. Anai, and Y. Takagi. 1998. Inter-locus relationship between genes controlling palmitic acid contents in soybean mutants. Breed. Sci. 48:377–381.

Li, Z., R.F. Wilson, W.E. Rayford, and H.R. Boerma. 2002. Molecular mapping genes conditioning reduced palmitic acid content in N87-2122-4 soybean. Crop Sci. 42:373–378.

Murray, M.G., and W.F. Thompson. 1980. Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. 8:4321–4325.

Ndzana, X., W.R. Fehr, G.A. Welke, E.G. Hammond, D.N. Duvick, and S.R. Cianzio. 1994. Influence of reduced palmitate content on agronomic and seed traits of soybean. Crop Sci. 34:646–649.

Nickell, A.D., J.R. Wilcox, and J.F. Cavins. 1991. Genetic relationships between loci controlling palmitic and linolenic acids in soybean. Crop Sci. 31:1169–1171.

Pfeil, B.E., J.A. Schlueter, R.C. Shoemaker, and J.J. Doyle. 2005. Placing paleopolyploidy in relation to taxon divergence: A phylogenetic analysis in legumes using 39 gene families. Syst. Biol. 54:441–454.

Primomo, V. 2000. Inheritance and stability of palmitic acid alleles in soybeans (Glycine max L Merr.). MS Thesis, Department of Plant Agriculture, University of Guelph, Guelph, ON, Canada.

Primomo, V.S., D.E. Falk, G.R. Ablett, J.W. Tanner, and I. Rajcan. 2002. Inheritance and interaction of low palmitic and low linolenic soybean. Crop Sci. 42:31–36.

Rahman, S.M., T. Kinoshita, T. Anai, and Y. Takagi. 1996. Genetic analysis of palmitic acid content using two soybean mutants. Breed. Sci. 46:343–347.

Rebetzke, G.J., J.W. Burton, T.E. Carter, and R.F. Wilson. 1998a. Changes in agronomic and seed characteristics with selection for reduced palmitic acid content in soybean. Crop Sci. 38:297–302.

Rebetzke, G.J., J.W. Burton, T.E. Carter, and R.F. Wilson. 1998b. Genetic variation for modifiers controlling reduced saturated fatty acid content in soybean. Crop Sci. 38:303–308.

Rebetzke, G.J., V.R. Pantalone, J.W. Burton, T.E. Carter, and R.F. Wilson. 2001. Genetic background and environmental influence palmitate content of soybean seed oil. Crop Sci. 41:1731–1736.

SAS Institute, Inc. 1996. SAS system for mixed models. SAS Institute Inc., Cary. NC.

SAS Institute, Inc. 1999–2001. SAS/STAT software: Release 8.2. SAS Institute Inc., Cary, NC.

Schlueter, J.A., P. Dixon, C. Granger, D. Grant, L. Clark, J.J. Doyle, and R.C. Shoemaker. 2004. Mining EST database to resolve evolutionary events in major crop species. Genome 47:868–876.

Schnebly, S.R., W.R. Fehr, G.A. Welke, E.G. Hammond, and D.N. Duvick. 1994. Inheritance of reduced and elevated palmitate in mutant lines of soybean. Crop Sci. 34:829–833.

Shoemaker, R.C., K. Polzin, J. Labate, J. Specht, E.C. Brummer, T. Olson, N. Young, V. Concibido, J. Wilcox, J.P. Tamulonis, G. Kochert, and H.R. Boerma. 1996. Genome duplication in soybean (*Glycine* subgenus *soja*). Genetics 144:329–338.

Stojšin, D., G.R. Ablett, B.M. Luzzi, and J.W. Tanner. 1998. Use of gene substitution values to quantify partial dominance in low palmitic acid soybean. Crop Sci. 38:1437–1441.

Thompson, J.D., D.G. Higgins, and T.J. Gibson. 1994. Clustal W–improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific GAP penalties and weight matrix choice. Nucleic Acids Res. 22:4673–4680.

Voelker, T.A. 1996. Plant acyl-ACP thioesterases: Chain-length determining enzymes in plant fatty acid biosynthesis. p. 111–133. *In* J.K. Setlow (ed.) Genetic engineering. Plenum, New York.

Wilcox, J.R., and J.F. Cavins. 1990. Registration of C1726 and C1727 soybean germplasm with altered levels of palmitic acid in seed oil of soybean. Crop Sci. 30:240.

Wilcox, J.R., J.W. Burton, G.J. Rebetzke, and R.F. Wilson. 1994. Transgressive segregation for palmitic acid in seed oil of soybean. 1994. Crop Sci. 34:1248–1250.

Wilson, R.F., J.W. Burton, W.P. Novitzky, and R.E. Dewey. 2001a. Current and future innovations in soybean (*Glycine max* L. Merr.) oil composition. J. Oleo Sci. 50:353–358.

Wilson, R.F., T.C. Marquardt, W.P. Novitzky, J.W. Burton, J.R. Wilcox, A.J. Kinney, and R.E. Dewey. 2001b. Metabolic mechanisms associated with alleles governing the 16:0 concentration of soybean oil. J. Am. Oil Chem. Soc. 78:335–340.

Wilson, R.F., T.C. Marquardt, W.P. Novitzky, J.W. Burton, J.R. Wilcox, and R.E. Dewey. 2001c. Effect of alleles governing 16:0 concentration on glycerolipid composition in developing soybeans. J. Am. Oil Chem. Soc. 78:329–334.

```
FATB1a    1 GACATAGTTCAAGTGGACACTTGGGTTTCTGGATCAGGGAAGAATGGTATGCGCCGTGATTGGCTTTTACGTGAC
FATB1b    1 GACGTAGTTCAAGTAGACACCTGGGCTTCTGGATCAGGGAAGAATGCTATGCGCCGTGATTGGGTTTTACGTGAC

FATB1a   76 TGCAAAACTGGTGAAATCTTGACAAGAGCTTCCAGgtagaaatcattctctggaattttccttccccttccttc
FATB1b   76 TGCAAAACTGGTGAAATCTTGACAAGAGCTTCCAGgtagaaatcattctctgtaattttccttccccttccttc

FATB1a  151 tgcttcaagcaaattttaagatgtgtatcttaatgtacttgatggtgattgggcacaattttgaatcttccatac
FATB1b  151 tgcttcaagcaaattttaagatgtgtatcttaatgtgcacgatgctgattggacacaattttaaatctttcaaac

FATB1a  226 attt.taaaagttatggaacccttttcttttccttctt.aagatgcaaatttgtcatgactgaagtttcaggtaat
FATB1b  226 atttacaaaagttatggaacccttttcttttctctcttgaagatgcaaatttgtcacgactgaagtttgaggaaat

FATB1a  299 catttgcattttgcagtgtt.aaaaagataatgaactacacatttattatattttgcaggcaaaaacctctaatt
FATB1b  301 catttgaattttgcaatgttaaaaaagataatgaactac.........atattttgcaggcaaaaacctctaatt

FATB1a  373 aaacaaactgaacattgtatcttagtttatttatcagactttatcatgtgtactgatgcatcaccttggagcttg
FATB1b  367 gaacaaactgaacattgtatcttagtttatttatcagactttatcatgtgtactgatgcatcaccttggagcttg

FATB1a  448 taatgaattacatattagtattttctgaactgtttgttatggttttggtgatctacagTGTTTGGGTCATGATGA
FATB1b  442 taatgaattacatattagcattttctgaactgtatgttatggttttggtgatctacagTGTTTGGGTCATGATGA

FATB1a  523 ATAAGCTAACACGGAGGCTGTCTAAAATTCCAGAAGAAGTCAGACAGGAGATAGGATCTTATTTTGTGGATTCTG
FATB1b  517 ATAAGCTGACACGGAGGCTGTCTAAAATTCCAGAAGAAGTCAGACAGGAGATAGGATCTTATTTTGTGGATTCTG

FATB1a  598 ATCCAATTCTGGAAGAGGATAACAGAAAACTGACTAAACTTGACGACAACACAGCGGATTATATTCGTACCGGTT
FATB1b  592 ATCCAATTCTAGAAGAGGATAACAGAAAACTGACTAAACTTGACGACAACACAGCAGATTATATTCGTACCGGTT

FATB1a  673 TAAgtgtatgtcaactagttttttttctaattgctgtcattaatttattttctcaaattatttcagatgttgttt
FATB1b  667 TAAgtgtatgtcaactag.ttttttttgtaattgttgtcattaatttcttttcttaaattatttcagatgttgctt

FATB1a  748 tctaattagtttacataatgcatcttcatt.ttgcagCCTAGGTGGAGTGATCTAGATATCAATCAGCATGTCAA
FATB1b  741 tctaattagtttacattatgtatcttcattcttccagTCTAGGTGGAGTGATCTAGATATCAATCAGCATGTCAA

FATB1a  822 CAATGTGAAGTACATTGGCTGGATTCTGGAGgtatttttctgttcttgtattctaatcaactgcaatccatgtta
FATB1b  816 CAATGTGAAGTACATTGACTGGATTCTGGAGgtatttttctgttcttgtattctaatccactgcagtccttgttt

FATB1a  897 gttctttaaccaaaggactgtcttttgattg.ttgcagAGTGCTCCACAGCCAATCTTGGAGAGTCATGAGCTTT
FATB1b  891 tgttgttaaccaaaggactgtcctttgattgtttgcagAGTGCTCCACAGCCAATCTTGGAGAGTCATGAGCTTT

FATB1a  971 CTTCCATGACTTTAGAGTATAGGAGAGAGTGTGGTAGGGACAGTGTGCTGGATTCCCTGACTGCTGTATCTGGGG
FATB1b  966 CTTCCGTGACTTTAGAGTATAGGAGGGAGTGTGGTAGGGACAGTGTGCTGGATTCCCTGACTGCTGTATCTGGGG

FATB1a 1046 CCGACATGGGCAATCTAGCTCACAGCGGGCATGTTGAGTGCAAGCATTTGCTTCGACTGGAAAATGGTGCTGAGA
FATB1b 1041 CCGACATGGGCAATCTAGCTCACAGTGGGCATGTTGAGTGCAAGCATTTGCTTCGACTCGAAAATGGTGCTGAGA

FATB1a 1121 TTGTGAGGGGCAGGACTGAGTGGAGGCCCAAACCTGTGAACAACTTTGGTGTTGTGAA
FATB1b 1116 TTGTGAGGGGCAGGACTGAGTGGAGGCCCAAACCTATGAACAACATTGGTGTTGTGAA
```

Supplemental Fig. S1. Genomic region of *GmFATB1a* and *GmFATB1b* showing the locations of gene-specific primer pairs. Sequences corresponding to exons are indicated in uppercase and intron sequences are shown in lowercase. Molecular marker FATB1a-312 is defined by primers corresponding to the sequences shown in bold type and FATB1b-314 is derived using primers corresponding to sequences that are both bolded and underlined. Marker FATB1a-411 is defined by primers corresponding to sequences in italicized type, and FATB1b-417 is generating using primers corresponding to sequences that are both italicized and underlined. Position 1 of the genomic sequences depicted corresponds to position 630 with reference to the start codons of the *GmFATB1a* and *GmFATB1b* cDNAs. The partial genomic sequences have been deposited in GenBank as Accession numbers DQ861997 (*GmFATB1a*) and DQ861998 (*GmFATB1b*).